Figure 4. The complete DNA sequence of the coding region of *Naegleria* gene TTK (including the TAA stop codon). The segment of the gene that encodes thiaminase I is underlined.
(SEQ ID NO. 1)

```
ATGTCCACTCAACCAAAGACACTCACTGTTGGTCTCTTCCCATATCTTCCTTCTTGGAATGAAAATGGCAACGAAGTTAA
ATTGATCAATTTGATCAAGGATGTTTTGCCAACTCAGGTTTCCGGATATAATATCGAATATACCGAATTTGATTGTTACA
GTGATGCTAGTCTTCAAAGTCTTCCAGATGTTTTCTCAACTGATAGCATTTTCCTTCCATATCTTGTTTCTTTGGGTGGT
GTCAAGAGTTTGGATGAATCATTGGTTCGTGGTGTTACTGGTGATTTGCATAGTTTTGTTTCCTCAAGTGCCTCTGTCAA
TGGTTCCGTTTATGGTTTCCCACAATACTTGTGCTCAAACTTTTTATTGTCCTCACCAAATGGTACTCAACAAGCATCTT
CCCTTTTAGAATTGGCTCAAAAGGTTGGTTATGAACAAATTGTTTATCCAGATGTTGCCTCTTCTAGTTCTTTCACAGTT
TTCGGATTGTATCAACAATTACTCCAATCATCATCATCAGCTGCAGTTGATATCAAGGCCTCTGATCTTCCACAATCTGG
TGACCAAGTCAACAAGGATATCACTCAAAAATATAGAACCATTTTGGATTCAACAGTTGTTGCCTCTCAAAGAGAATATA
TTAACTCTGTAAAGCAAGGTAAACCAATTTCAAACTACTATGTCGGATATAGTGAAAGTATGTGTGAAATTAAGGATATC
ATCAGAGATCAACAATACAATGTTCAACTCATTGGTACCTCTGATAAGCCATACGTTTATACTGATGTTTTGGCTTTGAA
TTCCAATTTGTGTGATGAAAAGCAAAAGGTTGCTGTTGAAGTTATCAAGAATTTATTGACTAATACTTTAGTTTTGGACT
TGTTGGGTCTCGGATTAACTCTCCCAGCCAACAAGAATGGTATTGCTCATTTGGCTAAATCATCAAACTTTTATGCTCAA
TTGAGCCAACAATTCGATGCCAAGGAAAGTGAAGTTAGAGTTTTGAGATGTGTTGACTTTGCTAACAAGGAAGTTAAGAA
TTGTGCTGGTGTCTTGAGACCATTCCTTCAACATATTGCTGTTGCTACTTTGCGTTGTTTGACTGCTGACACTGTCGAAA
AGGCTAAGAGTGGTCACCCTGGTATGCCAATTGGTATGTCACCAATTGCCTATGTTTTGTGGAAGTTCTTCTTCAAATCA
TCTAAGGATGATGTCAATTGGTTGAACAGAGATAGATTTGTTTTGAGTAATGGTCACGGTTGTACATTGCTTTATGCCAT
GTTGCACCTCACTGATTGTAACTTGAGTTTGGATGATCTCAAGAATTTCAGAAGTTTGCATTCCAAGACTCCTGGTCACC
CAGAATATGGTCACACTGAAGGTGTTGATGCTACTACTGGTCCATTGGGTCAAGGTGTTTGTAATGCTATTGGTATGGCT
CTCTCTGAAGCTCACTTGGCTGCTCGTTTCAATAAGGATGGACAAAATATCTTTGATCACCACACCTATGTTTTCCTTGG
TGATGGTTGTTTGATGGAACGTGTTGCTATGGAAGGTCTCTCATTTGCTGGTCACCAAAAGTTGAACAAGTTGATTGTTT
TCTATGATGACAATAGTATTACTATTGATGGTAAGACTGAATTGACCTTTACTCAAAATACTCCAGAAGTCATGAGAGGT
TTTGGATGGCACGTAATTGTTGTCGACAAGGCTGATAATGACTTGGTTGGTATTAAGGAAGCTATTTTGGAAGCTCACAC
TGTTACTGACAAGCCAATCATGATCGTTTGTAAGACTACAATTGGTTATTCCTCAAAGGTTCAAGGTACTGCTAAGGTTC
ACGGTTCTCCATTGGGTGCTGATGGATTGAAGAATTTGAAGGAAACTTGTGGTTTCACTGGTAATGATTTCTTCCATGTT
CCAGAAATTGTCAGAAAGGACTTTGCTACTGTCATTAATAGAAATAGTGAAAAGCTCTCTCAATGGAAGCAAGTTAAATC
TGCCTATGATACCACTCATGCTACTGAATCCCAACTCCTCCAAAGAATGATTAATCACGAATTGGAAGGTGATGTTATGG
AAAAGTTGCCAAAATACCTCGAACAAAAGAAGATTGCTACCAGATCTACATCTCAACAAGTTTTGAATGCCATCTATCCA
CTCATTCCTTCTCTCGTTGGTGGTTCAGCTGACTTGACTCCATCCAACTTGACTGATGTAACTGGATGTCAAGATTTCCA
ACCAAACAATAGAGTTGGTAGATATATCAGATTTGGTGTCCGTGAACATGCCATGGTTGCTATTGCCAATGGTATTCTCT
ATCATGGTGTTCTTAGAACCTATGTTGGTACATTCTTGAACTTTGCTTCATATGCTTTGGGTGCTATCAGATTGAGTGCC
TTGTCTGGTCTTCCAAATATTTATGTTTTCACTCATGACAGTATTGGTCTTGGTCAAGATGGTCCAACTCACCAACCTGT
TGAAGTTTTACCAATGTTGATAGCCATTCCAAATCACATTGTTTTCAGACCTGCTGATGGTAGAGAAACCAGTGGTGCTT
ATTTGTGGGCTGTTCAATCAAAGAAGACTCCATCCTCAATGATTCTTTCTCGTCAAGATTTGCCACAATTGACTGGTACT
GATATTTCAAAGGTTGCTTTGGGTGCCTATGTTATCCAAGGTGATGCTACTCCTGATGTTGTCCTTGTTGGTACTGGTTC
TGAAGTTTCCCTCATGGTTGAAGCTGCTGAAAAGTTGAAGGCTAACCTTAAGGTTAACGTTGTTTCCATGCCAAGTTGGG
AATTGTTTGTTCGTCAATCAGAAGAATACAGGAAGACTGTCTTCCCAGATGGTATTCCAGTTGTCAGTGCCGAAGCTTCA
TCAACCTTTGGTTGGACAAGCTTTGCTCACTATGCTGTTGGTATGACTACTTTCGGTGCTAGTGCTGCTGCTGAAGAAGT
TTACAAACTCCTCAAGATTACCTCAGACAATGTTGCTGAAAAGGCCACCAAATTGGTTACCAAGTATGGTAAGCAAGCTC
CAAGACTCAGCTTGTCTCTTGTTGGTGAAGAACTCTAA
```

Figure 5. The amino acid sequence encoded by the entire *Naegleria* TTK gene

(SEQ ID NO. 2)

```
MSTQPKTLTVGLFPYLPSWNENGNEVKLINLIKDVLPTQVSGYNIEYTEFDCYSDASLQSLPDVFSTDSIFLPYLVSLGG
VKSLDESLVRGVTGDLHSFVSSSASVNGSVYGFPQYLCSNFLLSSPNGTQQASSLLELAQKVGYEQIVYPDVASSSSFTV
FGLYQQLLQSSSSAAVDIKASDLPQSGDQVNKDITQKYRTILDSTVVASQREYINSVKQGKPISNYYVGYSESMCEIKDI
IRDQQYNVQLIGTSDKPYVYTDVLALNSNLCDEKQKVAVEVIKNLLTNTLVLDLLGLGLTLPANKNGIAHLAKSSNFYAQ
LSQQFDAKESEVRVLRCVDFANKEVKNCAGVLRPFLQHIAVATLRCLTADTVEKAKSGHPGMPIGMSPIAYVLWKFFFKS
SKDDVNWLNRDRFVLSNGHGCTLLYAMLHLTDCNLSLDDLKNFRSLHSKTPGHPEYGHTEGVDATTGPLGQGVCNAIGMA
LSEAHLAARFNKDGQNIFDHHTYVFLGDGCLMERVAMEGLSFAGHQKLNKLIVFYDDNSITIDGKTELTFTQNTPEVMRG
FGWHVIVVDKADNDLVGIKEAILEAHTVTDKPIMIVCKTTIGYSSKVQGTAKVHGSPLGADGLKNLKETCGFTGNDFFHV
PEIVRKDFATVINRNSEKLSQWKQVKSAYDTTHATESQLLQRMINHELEGDVMEKLPKYLEQKKIATRSTSQQVLNAIYP
LIPSLVGGSADLTPSNLTDVTGCQDFQPNNRVGRYIRFGVREHAMVAIANGILYHGVLRTYVGTFLNFASYALGAIRLSA
LSGLPNIYVFTHDSIGLGQDGPTHQPVEVLPMLIAIPNHIVFRPADGRETSGAYLWAVQSKKTPSSMILSRQDLPQLTGT
DISKVALGAYVIQGDATPDVVLVGTGSEVSLMVEAAEKLKANLKVNVVSMPSWELFVRQSEEYRKTVFPDGIPVVSAEAS
STFGWTSFAHYAVGMTTFGASAAAEEVYKLLKITSDNVAEKATKLVTKYGKQAPRLSLSLVGEEL
```

Figure 6. The DNA sequence of the 1068 bp *Naegleria* thiaminase I segment, as cloned in pNB1+ and expressed from the 3414 base genomic sequence of gene TTK.

(SEQ ID No. 3)

ATGTCCACTCAACCAAAGACACTCACTGTTGGTCTCTTCCCATATCTTCCTTCTTGGAATGAAAATGGCAACGAAGTTAA

ATTGATCAATTTGATCAAGGATGTTTTGCCAACTCAGGTTTCCGGATATAATATCGAATATACCGAATTTGATTGTTACA

GTGATGCTAGTCTTCAAAGTCTTCCAGATGTTTTCTCAACTGATAGCATTTTCCTTCCATATCTTGTTTCTTTGGGTGGT

GTCAAGAGTTTGGATGAATCATTGGTTCGTGGTGTTACTGGTGATTTGCATAGTTTTGTTTCCTCAAGTGCCTCTGTCAA

TGGTTCCGTTTATGGTTTCCCACAATACTTGTGCTCAAACTTTTTATTGTCCTCACCAAATGGTACTCAACAAGCATCTT

CCCTTTTAGAATTGGCTCAAAAGGTTGGTTATGAACAAATTGTTTATCCAGATGTTGCCTCTTCTAGTTCTTTCACAGTT

TTCGGATTGTATCAACAATTACTCCAATCATCATCATCAGCTGCAGTTGATATCAAGGCCTCTGATCTTCCACAATCTGG

TGACCAAGTCAACAAGGATATCACTCAAAAATATAGAACCATTTTGGATTCAACAGTTGTTGCCTCTCAAAGAGAATATA

TTAACTCTGTAAAGCAAGGTAAACCAATTTCAAACTACTATGTCGGATATAGTGAAAGTATGTGTGAAATTAAGGATATC

ATCAGAGATCAACAATACAATGTTCAACTCATTGGTACCTCTGATAAGCCATACGTTTATACTGATGTTTTGGCTTTGAA

TTCCAATTTGTGTGATGAAAAGCAAAAGGTTGCTGTTGAAGTTATCAAGAATTTATTGACTAATACTTTAGTTTTGGACT

TGTTGGGTCTCGGATTAACTCTCCCAGCCAACAAGAATGGTATTGCTCATTTGGCTAAATCATCAAACTTTTATGCTCAA

TTGAGCCAACAATTCGATGCCAAGGAAAGTGAAGTTAGAGTTTTGAGATGTGTTGACTTTGCTAACAAGGAAGTTAAGAA

TTGTGCTGGTGTCTTGAGACCATTCCTT

Figure 7. The encoded amino acid sequence of the expressed *Naegleria* thiaminase I gene, along with its codons, and with residues numbered (nucleotide/amino acid).
(SEQ ID NO. 3/SEQ ID NO. 4)

```
1/1                                  31/11
ATG TCC ACT CAA CCA AAG ACA CTC ACT GTT GGT CTC TTC CCA TAT CTT CCT TCT TGG AAT
Met ser thr gln pro lys thr leu thr val gly leu phe pro tyr leu pro ser trp asn


61/21                                91/31
GAA AAT GGC AAC GAA GTT AAA TTG ATC AAT TTG ATC AAG GAT GTT TTG CCA ACT CAG GTT
glu asn gly asn glu val lys leu ile asn leu ile lys asp val leu pro thr gln val


121/41                               151/51
TCC GGA TAT AAT ATC GAA TAT ACC GAA TTT GAT TGT TAC AGT GAT GCT AGT CTT CAA AGT
ser gly tyr asn ile glu tyr thr glu phe asp cys tyr ser asp ala ser leu gln ser


181/61                               211/71
CTT CCA GAT GTT TTC TCA ACT GAT AGC ATT TTC CTT CCA TAT CTT GTT TCT TTG GGT GGT
leu pro asp val phe ser thr asp ser ile phe leu pro tyr leu val ser leu gly gly


241/81                               271/91
GTC AAG AGT TTG GAT GAA TCA TTG GTT CGT GGT GTT ACT GGT GAT TTG CAT AGT TTT GTT
val lys ser leu asp glu ser leu val arg gly val thr gly asp leu his ser phe val


301/101                              331/111
TCC TCA AGT GCC TCT GTC AAT GGT TCC GTT TAT GGT TTC CCA CAA TAC TTG TGC TCA AAC
ser ser ser ala ser val asn gly ser val tyr gly phe pro gln tyr leu cys ser asn


361/121                              391/131
TTT TTA TTG TCC TCA CCA AAT GGT ACT CAA CAA GCA TCT TCC CTT TTA GAA TTG GCT CAA
phe leu leu ser ser pro asn gly thr gln gln ala ser ser leu leu glu leu ala gln


421/141                              451/151
AAG GTT GGT TAT GAA CAA ATT GTT TAT CCA GAT GTT GCC TCT TCT AGT TCT TTC ACA GTT
lys val gly tyr glu gln ile val tyr pro asp val ala ser ser ser ser phe thr val


481/161                              511/171
TTC GGA TTG TAT CAA CAA TTA CTC CAA TCA TCA TCA TCA GCT GCA GTT GAT ATC AAG GCC
phe gly leu tyr gln gln leu leu gln ser ser ser ser ala ala val asp ile lys ala


541/181                              571/191
TCT GAT CTT CCA CAA TCT GGT GAC CAA GTC AAC AAG GAT ATC ACT CAA AAA TAT AGA ACC
ser asp leu pro gln ser gly asp gln val asn lys asp ile thr gln lys tyr arg thr


601/201                              631/211
ATT TTG GAT TCA ACA GTT GTT GCC TCT CAA AGA GAA TAT ATT AAC TCT GTA AAG CAA GGT
ile leu asp ser thr val val ala ser gln arg glu tyr ile asn ser val lys gln gly


661/221                              691/231
AAA CCA ATT TCA AAC TAC TAT GTC GGA TAT AGT GAA AGT ATG TGT GAA ATT AAG GAT ATC
lys pro ile ser asn tyr tyr val gly tyr ser glu ser met cys glu ile lys asp ile


721/241                              751/251
ATC AGA GAT CAA CAA TAC AAT GTT CAA CTC ATT GGT ACC TCT GAT AAG CCA TAC GTT TAT
ile arg asp gln gln tyr asn val gln leu ile gly thr ser asp lys pro tyr val tyr


781/261                              811/271
ACT GAT GTT TTG GCT TTG AAT TCC AAT TTG TGT GAT GAA AAG CAA AAG GTT GCT GTT GAA
thr asp val leu ala leu asn ser asn leu cys asp glu lys gln lys val ala val glu


841/281                              871/291
GTT ATC AAG AAT TTA TTG ACT AAT ACT TTA GTT TTG GAC TTG TTG GGT CTC GGA TTA ACT
val ile lys asn leu leu thr asn thr leu val leu asp leu leu gly leu gly leu thr l


901/301                              931/311
CTC CCA GCC AAC AAG AAT GGT ATT GCT CAT TTG GCT AAA TCA TCA AAC TTT TAT GCT CAA
leu pro ala asn lys asn gly ile ala his leu ala lys ser ser asn phe tyr ala gln


961/321                              991/331
TTG AGC CAA CAA TTC GAT GCC AAG GAA AGT GAA GTT AGA GTT TTG AGA TGT GTT GAC TTT
leu ser gln gln phe asp ala lys glu ser glu val arg val leu arg cys val asp phe


1021/341                             1051/351
GCT AAC AAG GAA GTT AAG AAT TGT GCT GGT GTC TTG AGA CCA TTC CTT
ala asn lys glu val lys asn cys ala gly val leu arg pro phe leu
```

Figure 8. Comparison of the amino acid sequence of *Naegleria* thiaminase 1 (N40KAT) aligned with the sequences of *Bacillus* thiaminase (BSTP) and segments of the transketolases of six organisms. The numbering system indicated here is for the yeast transketolase.

Abbrev.		Genbank	Enzyme and Organism
ScTKT1ct	-	P23254	Transketolase of Saccharomyces cerevisiae
CpTKT7ct	-	Q42677	Transketolase of Craterostigma plantagineum
EcTKT2ct	-	P33570	Transketolase of Escherichia coli
BsTKTct	-	P45694	Transketolase of Bacillus subtilis
MgTKT	-	P47312	Transketolase of Mycoplasma genitalium
MjPTK1	-	Q58092	Transketolase of Methanococcus jannaschii
BSTP	-	P45741	Thiaminase I precursor from Bacillus thiaminolyticus
N40KAT	-	----->	Thiaminase I of Naegleria gruberi, aa 1-356

```
ScTKT1ct    ADDVK---QLKSKFGFNPDKSFVVPQEVYD-HYQKTILKPGVEANNKWNKLFSEYQKKFP    56
CpTKT7ct    PKEAE---ATRKNLGW-PYEPFHVPDDVKK-HWSRHIAE-GAALESAWNAKFAEFQKKFP
EcTKT2ct    EEEVA---LARQKLGW-HHPPFEIPKEIY--HAWDAREK-GEKAQQSWNEKFAAYKKAHP
BsTKTct     KEESK---LTKEAYAWTYEEDFYVPSEVYE-HFAVAVKESGEKKEQEWNAQFAKYKEVYP
MgTKT       EVDFQ---LFEKRTNT-NFNFFNYPDSIYH-WFKQTVIERQKQIKEDYNNLLISLKDK-P
MjPTK1      -----------------------------------------------------------
BsTP        MSKVKGFIYKPLMVMLALLLVVVSPAGAGAAHSDASSDITLKVAIYPYVPDPARFQAAVL
N40KAT      MSTQP----KTLTVGLFPYL----PS-----WNENGNEVKLINLIKDVLPT---------


ScTKT1ct    ELGAELARRLSGQLPANWESKLPTYTAKDSA----VATRKLSETVLEDVYNQLPELIGGS    112
CpTKT7ct    EEAADLKSIITGELPTNWESIFPTYTPENPG----LPTRTLSHQILNGLGDVLPGLLGGS
EcTKT2ct    QLAEEFTRRMSGGLPKDWEKTTQKYINELQANPAKIATRKASQNTLNAYGPMLPELLGGS
BsTKTct     ELAEQLELAIKGELPKDWDQEVPVYE-KGSS----LASRASSGEVLNGLAKKIPFFVGGS
MgTKT       -LFKKFTNWIDSDFQALYLNQLDEKKVAKKD----SATRNYLKDFLNQINNPNSNLYCLN
MjPTK1      --------------------MVKLSGVYKG------MRKGYGETLIELGKKYENLVVLD
BsTP        DQWQRQEPGVKLEF-TDWDSYSADPPDDLDV----FVLDSIFLSHFVDAGYLLP-FGSQD
N40KAT      -----QVSGYNIEY-TEFDCYSDASLQSLPD---VFSTDSIFLPYLVSLGGVKSLDESLV
                                                                  :


ScTKT1ct    ADLTPSNLTRWKEALDFQPPSSGSGNYSGRYIRYGIREHAMGA---IMNGISAFGANYKPYGG 172
CpTKT7ct    ADLTLSNMAFLKNSGDFQKKSPGE-----RNVKFGAREHAMGS---ICNGLALHSPGLLPYCA
EcTKT2ct    ADLAPSNLTIWKGSVSLKEDPAGN------YIHYGVREFGMTA---IANGIAHHG-GFVPYTA
BsTKTct     ADLAGSNKTTIKNAGDFTAVDYSG-----KNFWFGVREFAMGA---ALNGMALHG-GLRVFGG
MgTKT       ADVSRS--CFIKIGDDNLHENPCS-----RNIQIGIREFAMAT---IMNGMALHG-GIKVMGG
MjPTK1      ADLSGS-----TQTAMFAKEFPE------RFFNAGVAEQNMIG---MAAGLATTG--KIVFAS
BsTP        IDQAEDVLPFALQGAKRNGEVYGLP-------QILCTNLLFYRKGDLKIGQVDNIYELYKKIG
N40KAT      RGVTGDLHSFVSSSASVNGSVYGFP-------QYLCSNFLLSS----PNGTQQAS-SLLELAQ
               . :  .                        . .|            *
                                             Catalytic Cys


ScTKT1ct    TFLNFVS-YAAGAVR-LSALSGHPVIWVATHDSIGV-GE----DG-PTHQPIET--LAHF 222
CpTKT7ct    TYFVFTD-YMRAAMR-ISALSKARVLYIMTHDSIGL-GE----DG-PTHQPVEH--LASF
EcTKT2ct    TFLMFVE-YARNAAR-MAALMKARQIMVYTHDSIGL-GE----DG-PTHQAVEQ--LASL
BsTKTct     TFFVFSD-YLRPAIR-LAALMGLPVTYVFTHDSIAV-GE----DG-PTHEPVEQ--LASL
MgTKT       TFLAFAD-YSKPAIR-LGALMNLPVFYVYTHDSYQV-GG----DG-PTHQPYDQ--LPML
MjPTK1      SFSMFASGRAWEIIRNLVAYPKLNVKIVATHAGITV-GE----DG-ASHQMCED--IAIM
BsTP        TSHSEQIPPPQNKGLLINMAGGTTKASMYLEALIDVTGQYTEYDLLPPLDPLNDKVIRGL
N40KAT      KVGYEQIVYPDVASSSSFTVFGLYQQLLQSSSSAAV------DIKASDLPQSGD-QVNK
                            :              :          *  ..  .
```

(SEQ ID No.12)Peptide A -> ASDLPQSGD-QVNK

Figure 8, continued.

```
ScTKT1ct   RSLPNIQVWRPADGN-EVSAAYKNSLESKHTPSIIALSRQNLPQLEGS---SIESASKGG 278
CpTKT7ct   RAMPNILTLRPADGN-ETAGAYRAAVQNGERPSILVLARQKLPQLPGT---SIEGVSKGG
EcTKT2ct   RLTPNFSTWRPCDQV-EAAVGWKLAVERHNGPTALILSRQNLAQVERTPD-QVKEIARGG
BsTKTct    RAMPNLSLIRPADGN-ETAAAWKLAVQSTDHPTALVLTRQNLPTIDQTSEEALAGVEKGA
MgTKT      RAIENVCVFRPCDEK-ETCAGFNYGLLSQDQTTVLVLTRQPLKSIDNTD--SLKTL-KGG
MjPTK1     RAIPNMVVIAPTDYY-HTKNVIRTIAEYKG-PVYVRMPRRDTEIIYENEEEATFEIGKGK
BsTP       RLLINMAGEKPSQYVPEDGDAYVRASWFAQ-GSGRAFIGYSESMMRMG---DYAEQVRFK
N40KAT     DITQKYRTILDSTVV-ASQREYINSVKQGK-PISNYYVGYSESMCEIK---DIIRDQQYN
                      :                                               :
```

Peptide B -> TILDSTVV-ASQR (SEQ ID NO. 13)

```
ScTKT1ct   YVLQDVAN-----PDIILVATGSEVSLSVEAAKTLAAKNIKARVVSLPDFFTFDKQPLE- 332
CpTKT7ct   YVISDNSRGGNSKPDVILIGTGSELEIAARAGDELRKEGKKVRVVSLVCWELFAEQSEK-
EcTKT2ct   YVLKDSGG----KPDIILIATGSEMEITLQAAEKLAGEGRNVRVVSLPSTDIFDAQDEE-
BsTKTct    YVVSKSKN--E-TPDALLIASGSEVGLAIEAQAELAKENIDVSVVSMPSMDRFEKQSDE-
MgTKT      YILLDRKQ-----PDLIIAASGSEVQLAIEFEKVLTKQNVKVRILSVPNITLLLKQDEK-
MjPTK1     ILVDG--------EDLTIIATGEEVPEALRAGEILKENGISAEIVEMATIKPIDEEIIKK
BsTP       PISSSAG------QDIPLFYSDVVSVNSKTAHPELAKKLANVMASADTVEQALRPQADGQ
N40KAT     VQLIGTS-------DKPYVYTDVLALNSN---LCDEKQKVAVEVIKNLLTNTLVLDLLG-
                     *     :.       :           :      .      :  :
```

```
ScTKT1ct   YRLSVLPDNVPI-MSVEVLATTCWGKYAH------QSFGIDRFGASGKAPEVFKFFGFTP 385
CpTKT7ct   YRETVLPSGVTARVSVEAGSTFGWERFIGP-KG--KAVGIDRFGASAPAERLFKEFGITV
EcTKT2ct   YRESVLPSNVAARVAVEAGIADYWYKYVGL-KG--AIVGMTGYGESAPADKLFPFFGFTA
BsTKTct    YKNEVLPADVKKRLAIEMGSSFGWGKYTGL-EG--DVLGIDRFGASAPGETIINEYGFSV
MgTKT      YLKSLFDANSSL-ITIEASSSYEWFCFKKY-VKNHAHLGAFSFGESDDGDKVYQQKGFNL
MjPTK1     SKDFVVTVEDHSIIGGLGGAVAEVIASNGLNKKLLRIGINDVFGRSGKADELLKYYGLDG
BsTP       YPQYLLPARHQV-YEALMQDYPIYSELAQIVNK--PSNRVFRLGPEVRT-WLKDAKQVLP
N40KAT     -LGLTLPANKNG-IAHLAKSSNFYAQLSQ---------QFDAKESEVRVLRCVDFANKEV
```

Peptide C -> SSNFYAQLSQ---------QFDAK (SEQ ID NO. 14)

```
ScTKT1ct   EGVAERAQKTIAFYKGDKLISPLKKAF 412     (SEQ ID NO. 5)
CpTKT7ct   EAVVA-AAKEIC---------------         (SEQ ID NO. 6)
EcTKT2ct   ENIVAKAHKVLGVKGA-----------         (SEQ ID NO. 7)
BsTKTct    PNVVNRVKALINK--------------         (SEQ ID NO. 8)
MgTKT      ERLMKIFTSLRN---------------         (SEQ ID NO. 9)
MjPTK1     ESIAKRIMEEMKKE-------------         (SEQ ID NO. 10)
BsTP       EALGLTDVSSLAS--------------         (SEQ ID NO. 11)
N40KAT     KNCAGVLRPFL----------------         (SEQ ID NO. 4)
```